Improving Convergence Rate of Distributed Consensus Through Asymmetric Weights

He Hao, Prabir Barooah

Abstract

We propose a weight design method to increase the convergence rate of distributed consensus. Prior work has focused on symmetric weight design due to computational tractability. We show that with proper choice of *asymmetric* weights, the convergence rate can be improved significantly over even the symmetric optimal design. In particular, we prove that the convergence rate in a lattice graph can be made independent of the size of the graph with asymmetric weights. We then use a Sturm-Liouville operator to approximate the graph Laplacian of more general graphs. A general weight design method is proposed based on this continuum approximation. Numerical computations show that the resulting convergence rate with asymmetric weight design is improved considerably over that with symmetric optimal weights and Metropolis-Hastings weights.

I. INTRODUCTION

In distributed consensus, each agent in a network updates its state by aggregating the information from its neighbors so that all the agents' states reach a common value. Distributed consensus has been widely studied in recent times due to its wide ranging applications such as multi-vehicle rendezvous, data fusion in large sensor network, coordinated control of multi-agent system and formation flight of unmanned vehicles and clustered satellites, etc. (see [1]–[5] and references therein).

The topic of this paper is the convergence rate of distributed consensus protocols in graphs with fixed (time invariant) topology. The convergence rate is extremely important; it determines practical applicability of the protocol. If the convergence rate is small, it will take many iterations

He Hao and Prabir Barooah are with Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL 32611, USA, email: `hehao,pbarooah@ufl.edu`. This work was supported by the National Science Foundation through Grant CNS-0931885 and ECCS-0925534.

before the states of all agents are sufficiently close. Compared to the vast literature on design of consensus protocols, however, the literature on convergence rate analysis is meager. Convergence rate of distributed consensus in time-varying graphs have been studied in [6]–[8]. The related problem of mixing time of Markov chains is studied in [9]. In [10], convergence rate for a specific class of graphs, that we call L-Z geometric graphs, are established as a function of the number of agents. In general, the convergence rate of consensus algorithms tend to be slow, and decreases as the number of agents increases. It is shown in [11] that the convergence rate can be arbitrarily fast in small-world networks. However, networks in which communication is only possible between agents that are close enough are not likely to be small-world.

One of the seminal works on this subject is convex optimization of weights on edges of the graph to maximize the consensus convergence rate [12], [13]. Convex optimization imposes the constraint that the weights of the graph must be symmetric, which means any two neighboring agents put equal weight on the information received from each other. The convergence rate of consensus protocols on graphs with symmetric weights degrades considerably as the number of agents in the network increases. In a D-dimensional lattice, for instance, the convergence rate is $O(1/N^{2/D})$ if the weights are symmetric, where N is the number of agents. This result follows as a special case of the results in [10]. Thus, the convergence rate becomes arbitrarily small if the size of the network grows without bound.

In [14]–[16], finite-time distributed consensus protocols are proposed to improve the performance over asymptotic consensus. However, in general, the finite time needed to achieve consensus depends the number of agents in the network. Thus, for large size of networks, although consensus can be achieved in finite time, the time needed to reach consensus becomes large.

In this paper, we study the problem of how to increase the convergence rate of consensus protocols by designing *asymmetric* weights on edges. We first consider lattice graphs and derive precise formulae for convergence rate in these graphs. In particular, we show that in lattice graphs, with proper choice of asymmetric weights, the convergence rate of distributed consensus can be bounded away from zero uniformly in N. Thus, the proposed asymmetric design makes distributed consensus highly scalable. In addition, we provide exact formulae for asymptotic steady-state consensus value. With asymmetric weights, the consensus value in general is not the average of the initial conditions.

We next propose a weight design scheme for arbitrary 2-dimensional geometric graphs, i.e., graphs consisting of nodes in $\mathbb{R}^2$. Here we use the idea of continuum approximation to extend the asymmetric design from lattices to geometric graphs. We show how a Sturm-Liouville operator can be used to approximate the graph Laplacian in the case of lattices. The spectrum of the Laplacian and the convergence rate of consensus protocols are intimately related. The discrete weights in lattices can be seen as samples of a continuous weight function that appears in the S-L operator. Based on this analogy, a weight design algorithm is proposed in which a node i chooses the weight on the edge to a neighbor j depending on the relative angle between i and j. Numerical simulations show that the convergence rate with asymmetric designed weights in large graphs is an order of magnitude higher than that with (i) optimal symmetric weights, which are obtained by convex optimization [12], [13], and (ii) asymmetric weights obtained by Metropolis-Hastings method, which assigns weights uniformly to each edge connecting itself to its neighbor. The proposed weight design method is decentralized, every node can obtain its own weight based on the angular position measurements with its neighbors. In addition, it is computationally much cheaper than obtaining the optimal symmetric weights using convex optimization method. The proposed weight design method can be extended to geometric graphs in $\mathbb{R}^D$, but in this paper we limit ourselves to $\mathbb{R}^2$.

The rest of this paper is organized as follows. Section II presents the problem statement. Results on size-independent convergence rate on lattice graphs with asymmetric weight are stated in Section III. Asymmetric weight design method for more general graphs appear in Section IV. The paper ends with conclusions and future work in Section V.

II. PROBLEM STATEMENT

To study the problem of distributed linear consensus in networks, we first introduce some terminologies. The network of N agents is modeled by a graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ with vertex set $\mathbf{V} = \{1, \ldots, N\}$ and edge set $\mathbf{E} \subset \mathbf{V} \times \mathbf{V}$. We use (i, j) to represent a directed edge from i to j. A node i can receive information from j if and only if $(i, j) \in \mathbf{E}$. In this paper, we assume that communication is bidirectional, i.e. $(i, j) \in \mathbf{E}$ if and only if $(j, i) \in \mathbf{E}$. For each edge $(i, j) \in \mathbf{E}$ in the graph, we associate a weight $W_{i,j} > 0$ to it. The set of neighbors of i is defined as $\mathcal{N}_i := \{j \in \mathbf{V} : (i, j) \in \mathbf{E}\}$. The Laplacian matrix L of an arbitrary graph $\mathbf{G}$ with

edge weights $W_{i,j}$ is defined as

$$
L_{i,j} = \begin{cases} -W_{i,j} & i \neq j, (i,j) \in \mathbf{E}, \\ \sum_{k=1}^{N} W_{i,k} & i = j, (i,k) \in \mathbf{E}, \\ 0 & \text{otherwise.} \end{cases} \tag{1}
$$

A linear consensus protocol is an iterative update law:

$$
x_i(k+1) = W_{i,i}\, x_i(k) + \sum_{j \in \mathcal{N}_i} W_{i,j}\, x_j(k), \quad i \in \mathbf{V}, \tag{2}
$$

with initial conditions $x_i(0) \in \mathbb{R}$, where $k = \{0,1,2,\cdots\}$ is the discrete time index. Following standard practice we assume the weight matrix W is a stochastic matrix, i.e. $W_{i,j} \geq 0$ and $W\mathbf{1} = \mathbf{1}$, where $\mathbf{1}$ is a vector with all entries of 1. The distributed consensus protocol (2) can be written in the following compact form:

$$
x(k+1) = Wx(k), \tag{3}
$$

where $x(k) = [x_1(k), x_2(k), \cdots, x_N(k)]^T$ is the states of the N agents at time k. It's straightforward to obtain the following relation $L = I - W$, where I is the $N \times N$ identity matrix and L is the Laplacian matrix associated with the graph with $W_{i,j}$ as its weights on the directed edge (i,j). In addition, their spectra are related by $\sigma(L) = 1 - \sigma(W)$, i.e. $\mu_\ell(L) = 1 - \lambda_\ell(W)$, where $\ell \in \{1,2,\cdots,N\}$ and μ_ℓ, λ_ℓ are the eigenvalues of L and W respectively. The linear distributed consensus protocol (3) implies $x(k) = W^k x(0)$. We assume W is strong connected (irreducible) and primitive. In that case the spectral radius of W is 1 and there is exactly one eigenvalue on the unit disk. Let $\pi \in \mathbb{R}^{1 \times N}$ be the left Perron vector of W corresponding to the eigenvalue of 1, i.e. $\pi W = \pi$, $\pi_i > 0$ and $\sum_{i=1}^{N} \pi_i = 1$, we have

$$
\lim_{k \to \infty} W^k = \mathbf{1}\pi, \tag{4}
$$

Therefore, all the states of the N agents asymptotically converge to a steady state value $\bar{x}$ as $k \to \infty$,

$$
\lim_{k \to \infty} x(k) = \mathbf{1}\pi x(0) = \mathbf{1}\bar{x}, \tag{5}
$$

where $\bar{x} = \sum_{i=1}^{N} \pi_i x_i(0)$.

One of the most important feature of linear distributed consensus is the rate of convergence to its steady state value. It's well known that for a primitive stochastic matrix, the rate of

convergence R can be measured by the spectral gap $R = 1 - \rho(W)$, where $\rho(W)$ is the essential spectral radius of W, which is defined as

$$\rho(W) := \max\{|\lambda| : \lambda \in \sigma(W) \setminus \{1\}\}.$$

If the eigenvalues of W are real and they are ordered in a non-increasing fashion such that $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$, then the convergence rate of W is given by

$$R = 1 - \rho(W) = \min\{1 - \lambda_2, 1 + \lambda_N\}. \tag{6}$$

In addition, from Gerschgorin circle theorem, we have that $\lambda_N \geq -1 + 2 \max_i W_{ii}$. If $\max_i W_{ii} \neq 0$, then $1 + \lambda_N$ is a constant bounded away from 0. Therefore, the key to find a lower bound for the convergence rate of W is to find an upper bound on the second largest eigenvalue λ_2 of W. Equivalently, we can find a lower bound of the second smallest eigenvalue μ_2 of the associated Laplacian matrix L, since $\mu_2 = 1 - \lambda_2$.

Definition 1: We say a graph **G** has *symmetric* weights if $W_{i,j} = W_{j,i}$ for each pair of neighboring agents $(i,j) \in \mathbf{E}$. Otherwise, the weights are called *asymmetric*. □ If the weights are symmetric, the matrix W is doubly stochastic, meaning that each row and column sum is 1.

The following theorem summaries the results in [10] on the convergence rate of consensus with symmetric weights in a broad class of graphs that include lattices. A D-dimensional lattice, specifically a $N_1 \times N_2 \times \cdots \times N_D$ lattice, is a graph with $N = N_1 \times N_2 \times \cdots \times N_D$ nodes, in which the nodes are placed at the integer unit coordinate points of the D-dimensional Euclidean space and each node connects to other nodes that are exactly one unit away from it. A D-dimensional lattice is drawn in $\mathbb{R}^D$ with a Cartesian reference frame whose axes are denoted by $x_1, x_2, \cdots, x_D$. We call a graph is a *L-Z geometric graph* if it can be seen as a perturbation of regular lattice in D-dimensional space; each node connects other nodes within a certain range. The formal definition is given in [10].

Theorem 1: Let **G** be a D-dimensional connected L-Z geometric graph or lattice and let W be any doubly stochastic matrix compatible with **G**. Then

$$\frac{c_1}{N^{2/D}} \leq R \leq \frac{c_2}{N^{2/D}}, \tag{7}$$

where N is the number of nodes in the graph **G** and c_1, c_2 are some constants independent of N. □

The above theorem states that for any connected L-Z geometric/lattice graph $\mathbf{G}$, the convergence rate of consensus with symmetric weights cannot be bounded away from 0 uniformly with the size N of the graph. The convergence rate of the network becomes arbitrarily slow as N increases without bound. The loss of convergence rate with symmetric information graph has also been observed in the vehicular formation [17], [18]. In fact, another important conclusion of the result above is that heterogeneity in weights among nodes, as long as W is symmetric, does not change the asymptotic scaling of the convergence rate. At best it can change the constant in front of the scaling formula (see [9] also). Therefore, even centralized weight optimization scheme proposed in [12], [13] - that constrain the eights to be symmetric in order to make the optimization problem convex - will suffer from the same issue as that of un-optimized weights on the edges. Namely, the convergence rate will decay as $O(1/N^{2/D})$ in a D-dimensional lattice/L-Z geometric graph even with the optimized weights. In the rest of the paper, we study the problem of speeding up the convergence rate by designing *asymmetric* weights.

III. FAST CONSENSUS ON D-DIMENSIONAL LATTICES

First we establish technical results (whose proofs are provided in the appendix) on the spectrum and Perron vectors of D-dimensional lattices with asymmetric weights on the edges. We then summarize their design implications at the end of section III-A.

A. Asymmetric weights in lattices

We first consider distributed consensus on a 1-dimensional lattice. This will be useful in generalizing to D-dimensional lattices. Each agent interacts with its nearest neighbors in the lattice (one on each side). Its information graph is depicted in Figure 1. The updating law of agent i is given by

$$x_i(k+1) = W_{i,i}x_i(k) + W_{i,i-1}x_{i-1}(k) + W_{i,i+1}x_{i+1}(k).$$

where $i \in \{2, 3, \cdots, N-1\}$. The updating laws of the 1-st and N-th agents are slightly different from the above equation, since they only have one neighbor.
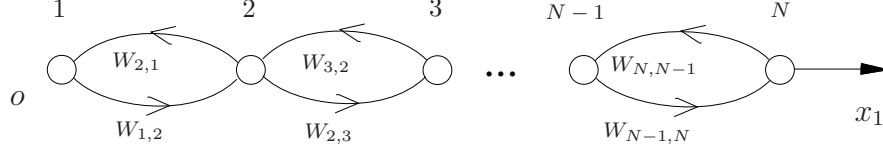
Fig. 1.   Information graph for a 1-D lattice of N agents.

The weight matrix $W^{(1)}$ for the 1-dimensional lattice is tridiagonal:

$$W^{(1)} = \begin{bmatrix} W_{1,1} & W_{1,2} & & & \\ W_{2,1} & W_{2,2} & W_{2,3} & & \\ & \ddots & & \ddots & \\ & W_{N-1,N-2} & W_{N-1,N-1} & W_{N-1,N} \\ & & W_{N,N-1} & W_{N,N} \end{bmatrix}.$$

The following lemma gives the spectrum and the left-hand Perron vector for the weight matrix $W^{(1)}$.

Lemma 1: Let $W^{(1)}$ be the weight matrix associated with the 1-dimensional lattice with the weights given by $W_{i,i+1} = c, W_{i+1,i} = a$, where $a \neq c$ are positive constants and $a + c \leq 1$. Then its eigenvalue are

$$\lambda_1 = 1, \quad \lambda_\ell = 1 - a - c + 2\sqrt{ac} \cos \frac{(\ell - 1)\pi}{N},$$

where $\ell \in \{2, \cdots, N\}$, and its left Perron vector is

$$\pi = \frac{1 - c/a}{1 - (c/a)^N}[1, c/a, (c/a)^2, \cdots, (c/a)^{N-1}]. \qquad \square$$

We next consider consensus on a D-dimensional lattice with the following weights

$$W_{i,i^{d+}} = c_d, \quad W_{i,i^{d-}} = a_d, \tag{8}$$

where $a_d \neq c_d$ are positive constants and $\sum_{d=1}^{D} a_d + c_d \leq 1$. The notation i^{d+} denotes the neighbor on the positive x_d axis of node i and i^{d-} denotes the neighbor on the negative x_d axis of node i. For example, 2^{1+} and 2^{1-} in Figure 2 denote node 3 and node 1, respectively, and 2^{2+} is node 5.
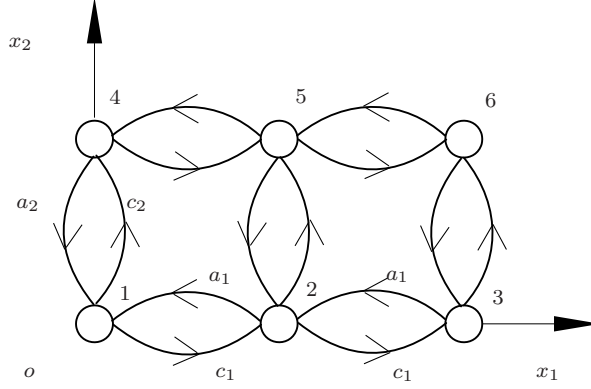
Fig. 2. A pictorial representation of a 2-dimensional lattice information graph with the weights $W_{i,i^{d+}}^{(2)} = c_d, W_{i,i^{d-}}^{(2)} = a_d$, where $d = 1, 2$.

Lemma 2: Let $W^{(D)}$ be the weight matrix associated with the D-dimensional lattice with the weights given in (8). Then its eigenvalues are given by

$$\lambda_{\vec{\ell}}\left(W^{(D)}\right) = 1 - \sum_{d=1}^{D}(1 - \lambda_{\ell_d}(W_d^{(1)})),$$

where $\vec{\ell} = (\ell_1, \ell_2, \cdots, \ell_D)$, in which $\ell_d \in \{1, 2, \cdots, N_d\}$ and $W_d^{(1)}$ is the $N_d \times N_d$ weight matrix associated with a 1-dimensional lattice with the weights given by $W_d^{(1)}(i, i+1) = c_d, W_d^{(1)}(i+1, i) = a_d$ and $i \in \{1, \cdots, N_d - 1\}$. Its left Perron vector is $\pi = \pi_D^{(1)} \otimes \pi_{D-1}^{(1)} \otimes \cdots \otimes \pi_1^{(1)}$, where $\pi_d^{(1)}$ is the left Perron vector of $W_d^{(1)}$. $\square$

The next theorem shows the implications of the preceding technical results on the convergence rate in D-dimensional lattices.

Theorem 2: Let **G** be a D-dimensional lattice graph and let $W^{(D)}$ be an asymmetric stochastic matrix compatible with **G** with the weights given in (8). Then the convergence rate satisfies

$$R \geq c_0, \tag{9}$$

where $c_0 \in (0, 1)$ is a constant independent of N. $\square$

Remark 1: Recall from Theorem 1, for any L-Z geometric or lattice graphs, as long as the weight matrix W is symmetric, no matter how do we design the weights $W_{i,j}$, the convergence rate becomes progressively smaller as the number of agents N increases, and it cannot be uniformly bounded away from 0. In contrast, Theorem 2 shows that for lattice graphs, asymmetry

in the weights makes the convergence rate uniformly bounded away from 0. In fact, any amount of asymmetry along the coordinate axes of the lattice ($a_d \neq c_d$), will make this happen. Asymmetric weights thus make the linear distributed consensus law highly scalable. It eliminates the problem of degeneration of convergence rate with increasing N.

The second question is where do the node states converge to with asymmetric weights? Recall that the asymptotic steady state value of all agents is $\bar{x} = \sum_{i=1}^{N} \pi_i x_i(0)$. For a lattice graph, its Perron vector π is given in Lemma 1 and Lemma 2. Thus we can determine the steady state value $\bar{x}$ if the initial value $x(0)$ is given. This information is particularly useful to find the rendezvous position in multi-vehicle rendezvous problem. On the other hand, we see from Lemma 1 and Lemma 2 that if $a_d \neq c_d$, then $\pi_i \neq \frac{1}{N}$, which implies the steady-state value is not the average of the initial values. The asymmetric weight design is not applicable to distributed averaging problem. $\qquad\square$

B. Numerical comparison

In this section, we present the numerical comparison of the convergence rates of the distributed protocol (3) between asymmetric designed weights (Theorem 2) and symmetric optimal weights obtained from convex optimization [12], [13]. For simplicity, we take the 1-D lattice as an example. The asymmetric weights used are $W_{i,i+1} = c = 0.3, W_{i+1,i} = a = 0.2$. We see from Figure 3 that the convergence rate with asymmetric designed weights is much larger than that with symmetric optimal weights. In addition, given the asymmetric weight values $c = 0.3, a = 0.2$, we obtain from Lemma 1 that $\lambda_2 \leq 0.5 + 2\sqrt{0.06}, \lambda_N \geq 0.5 + 2\sqrt{0.06}$, which implies

$$R = \min\{1 - \lambda_2, 1 + \lambda_N\} \geq 0.5 - 2\sqrt{0.06}. \tag{10}$$

We see from Figure 3 that the convergence rate R is indeed uniformly bounded below by (10).

IV. FAST CONSENSUS IN MORE GENERAL GRAPHS

In this section, we study how to design the weight matrix W to increase the convergence rate of consensus in graphs that are more general than lattices. We use the idea of continuum approximation. Under some "niceness" properties, a graph can be thought of as approximation of a D-dimensional lattice, and by extension, of the Euclidean space corresponding to $\mathbb{R}^D$ [19].
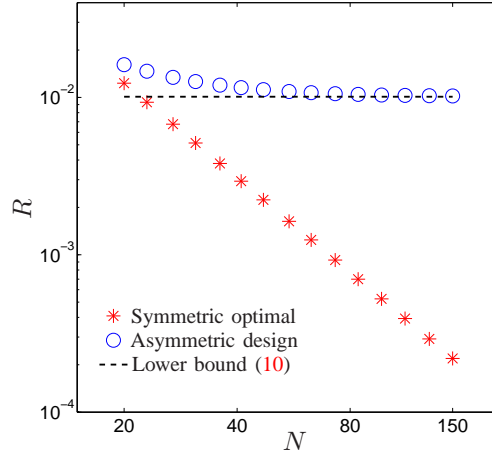
Fig. 3. Comparison of convergence rate of 1-D lattice between asymmetric design and convex optimization (symmetric optimal).

These properties have to do with the graph not having arbitrarily large holes etc. Precise conditions under which a graph can be approximated by the D-dimensional lattice are explored in [20] (for infinite graphs) and in [10] (for finite graphs). The dimension D of the corresponding lattice/Euclidean space is also determined by these properties.

The key is to embed the discrete graph problem into a continuum-domain problem. We use a Sturm-Liouville operator to approximate the Laplacian matrix of a D-dimensional geometric graph. A D-dimensional geometric graph is simply a graph with a mapping of nodes to points in $\mathbb{R}^D$. Based on this approximation, we re-derive the asymmetric weights for lattices described in the previous section as values of continuous functions defined over $\mathbb{R}^D$ along the principal axes in $\mathbb{R}^D$. In a lattice, the neighbors of a node lie along the principal canonical axes of $\mathbb{R}^D$. For an arbitrary graph, the weights are now chosen as samples of the same functions, along directions in which the neighbors lie.

The method is applicable to arbitrary dimension, but we only consider the 2-D case in this paper. Graphs with 2-D drawings are one of the most relevant classes of graphs for sensor networks where consensus is likely to find application.
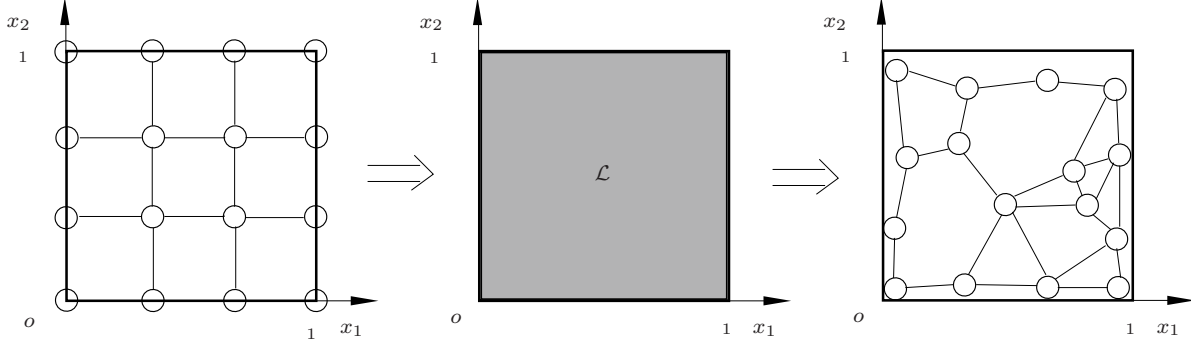
Fig. 4.   Continuum approximation of general graphs.

A. Continuum approximation

Recall that the convergence rate is intimately connected to the Laplacian matrix. We will show that the Laplacian matrix associated with a large 2-D lattice with certain weights can be approximated by a Sturm-Liouville operator defined on a 2-D plane. Thus it's reasonable to suppose that the Sturm-Liouville operator is also a good (continuum) approximation of the Laplacian matrix of large graphs with 2-D drawing. We start from 2-D lattice graph and derive a Sturm-Liouville operator. We then use this operator to approximate the graph Laplacian of more general graphs. The idea is illustrated in Figure 4.

For ease of description, we first consider a 1-D lattice, with the following asymmetric weights inspired by [21],

$$W_{i,i+1} = c = \frac{1+\varepsilon}{2}, \quad W_{i+1,i} = a = \frac{1-\varepsilon}{2}, \tag{11}$$

where $i \in \{1, 2, \cdots, N-1\}$ and $\varepsilon \in (0,1)$ is a constant. The graph Laplacian corresponding to the weights given in (11) is given by

$$L^{(1)} = \begin{bmatrix} \frac{1+\varepsilon}{2} & \frac{-1-\varepsilon}{2} & & & \\ \frac{-1+\varepsilon}{2} & 1 & \frac{-1-\varepsilon}{2} & & \\ & \ddots & \ddots & \ddots & \\ & & \frac{-1+\varepsilon}{2} & 1 & \frac{-1-\varepsilon}{2} \\ & & & \frac{-1+\varepsilon}{2} & \frac{1-\varepsilon}{2} \end{bmatrix}. \tag{12}$$

Recall that to find a lower bound of the convergence rate of the weight matrix $W^{(1)}$, it's sufficient to find a lower bound of the second smallest eigenvalue of the associate Laplacian matrix $L^{(1)}$.

We now use a Sturm-Liouville operator to approximate the Laplacian matrix $L^{(1)}$. We first consider the finite-dimensional eigenvalue problem $L^{(1)}\phi = \mu\phi$,

$$
\begin{bmatrix}
\frac{1+\varepsilon}{2} & \frac{-1-\varepsilon}{2} & & & \\
\frac{-1+\varepsilon}{2} & 1 & \frac{-1-\varepsilon}{2} & & \\
& \ddots & \ddots & \ddots & \\
& & \frac{-1+\varepsilon}{2} & 1 & \frac{-1-\varepsilon}{2} \\
& & & \frac{-1+\varepsilon}{2} & \frac{1-\varepsilon}{2}
\end{bmatrix}
\begin{bmatrix}
\phi_1 \\ \phi_2 \\ \vdots \\ \phi_{N-1} \\ \phi_N
\end{bmatrix}
= \mu
\begin{bmatrix}
\phi_1 \\ \phi_2 \\ \vdots \\ \phi_{N-1} \\ \phi_N
\end{bmatrix}.
$$

Expanding the equation, we have the following coupled difference equations

$$
\frac{-1+\varepsilon}{2}\phi_{i-1} + \phi_i + \frac{-1-\varepsilon}{2}\phi_{i+1} = \mu\phi_i,
$$

where $i \in \{1, 2, \cdots, N\}$ and $\phi_0 = \phi_1$, $\phi_{N+1} = \phi_N$. The above equation can be rewritten as

$$
-\frac{1}{2N^2}\frac{\phi_{i-1} - 2\phi_i + \phi_{i+1}}{1/N^2} - \frac{\varepsilon}{N}\frac{\phi_{i+1} - \phi_{i-1}}{2/N} = \mu\phi_i.
$$

The starting point for the continuum approximation is to consider a function $\phi(x) : [0, 1] \to \mathbb{R}$ that satisfies:

$$
\phi_i = \phi(x)|_{x=i/(N+1)}, \tag{13}
$$

such that functions that are defined at discrete points i will be approximated by functions that are defined everywhere in $[0, 1]$. The original functions are thought of as samples of their continuous approximations. Under the assumption that N is large, using the following finite difference approximations:

$$
\left[\frac{\phi_{i-1} - 2\phi_i + \phi_{i+1}}{1/N^2}\right] = \left[\frac{\partial^2\phi(x,t)}{\partial x^2}\right]_{x=i/(N+1)},
$$

$$
\left[\frac{\phi_{i+1} - \phi_{i-1}}{2/N}\right] = \left[\frac{\partial\phi(x,t)}{\partial x}\right]_{x=i/(N+1)},
$$

the finite-dimensional eigenvalue problem can be approximated by the following Sturm-Liouville eigenvalue problem

$$
\mathcal{L}^{(1)}\phi(x) = \mu\phi(x), \quad \mathcal{L}^{(1)} = -\frac{1}{2N^2}\frac{d^2}{dx^2} - \frac{\varepsilon}{N}\frac{d}{dx}, \tag{14}
$$

with the following Neumann boundary conditions

$$
\frac{d\phi(0)}{dx} = \frac{d\phi(1)}{dx} = 0. \tag{15}
$$

Lemma 3: The eigenvalues of the Sturm-Liouville operator $\mathcal{L}^{(1)}$ (14) with boundary condition (15) for $0 < \varepsilon < 1$ are real and the first two smallest eigenvalues satisfy

$$\mu_1(\mathcal{L}^{(1)}) = 0, \quad \mu_2(\mathcal{L}^{(1)}) \geq \varepsilon^2/2. \qquad \qquad \square$$

We see from Lemma 3 that the second smallest eigenvalue of the Sturm-Liouville operator $\mathcal{L}^{(1)}$ is uniformly bounded away from zero. This result is not surprising, since it's a continuum counterpart of Lemma 1, which shows that the second smallest eigenvalue corresponding to the 1-D lattice with designed asymmetric weights is uniformly bounded below.

We now consider the distributed consensus on D-dimensional lattices. In particular, we consider the following weights on the graph

$$W_{i,i^{d+}}^{(D)} = c_d = \frac{1+\varepsilon}{2D}, \quad W_{i,i^{d-}}^{(D)} = a_d = \frac{1-\varepsilon}{2D}, \qquad (16)$$

where $\varepsilon \in (0, 1)$ is a constant.

The Laplacian matrix of the D-dimensional square lattices with the weights given in (16) is given by $L^{(D)} = I - W^{(D)}$. Following the similar procedure as the 1-dimensional lattice, the second smallest eigenvalue of the Laplacian matrix $L^{(D)}$ can be approximated by that of the following Sturm-Liouville operator

$$\mathcal{L}^{(D)} = -\sum_{\ell=1}^{D} (\frac{1}{2DN_d^2}\frac{d^2}{dx_d^2} + \frac{\varepsilon}{DN_d}\frac{d}{dx_d}), \qquad (17)$$

with the following Neumann boundary conditions

$$\left.\frac{\partial\phi(\vec{x})}{\partial x_d}\right|_{x_d=0 \text{ or } 1} = 0, \qquad (18)$$

where $d = 1, 2, \cdots, D$ and $\vec{x} = [x_1, x_2, \cdots, x_D]^T$.

The continuum approximation has been used to study the stability margin of large vehicular platoons [21], [22], in which the continuum model gives more insight on the effect of asymmetry on the stability margin of the systems. In this paper, we use the second smallest eigenvalue of the Sturm-Liouville operator $\mathcal{L}^{(D)}$ to approximate that of the Laplacian matrix $L^{(D)}$.

Theorem 3: The second smallest eigenvalues $\mu_2(\mathcal{L}^{(D)})$ of the Sturm-Liouville operator $\mathcal{L}^{(D)}$ (17) with boundary condition (18) for $0 < \varepsilon < 1$ is real and satisfies

$$\mu_2(\mathcal{L}^{(D)}) \geq \frac{\varepsilon^2}{2D}, \qquad (19)$$

which is a positive constant independent of N. $\qquad \qquad \square$

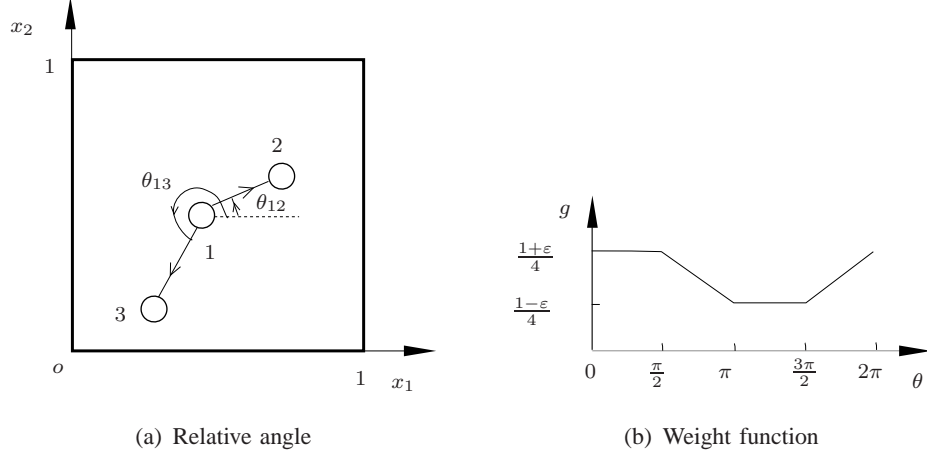(a) Relative angle   (b) Weight function

Fig. 5. Weight design for general graphs.

B. Weight design for general graphs

The inspiration of the proposed method comes from the design for lattices. The 4 weights for each node i in a 2-D lattice can be re-expressed as samples of a continuous function $g :$ $[0, 2\pi) \to [\frac{1-\epsilon}{4}, \frac{1+\epsilon}{4}]$:

$$W_{i,i^{1+}} = g(\theta_{i,i^{1+}}), \qquad\qquad W_{i,i^{2+}} = g(\theta_{i,i^{2+}}),$$

$$W_{i,i^{1-}} = g(\theta_{i,i^{1-}}), \qquad\qquad W_{i,i^{2-}} = g(\theta_{i,i^{2-}})$$

where $\theta_{i,j}$ is the relative angular position of j with respect to i. Given the angular positions of i's neighbors and the values of the weights, we know that the function g must satisfy:

$$g([0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}]) = [\frac{1+\varepsilon}{4}, \frac{1+\varepsilon}{4}, \frac{1-\varepsilon}{4}, \frac{1-\varepsilon}{4}]. \tag{20}$$

Thus, we choose the function g as shown in Figure 5 (b).

For an arbitrary graph, we now choose the weights by sampling the function according to the angle associated with each edge (i, j):

$$W_{i,k} = \frac{g(\theta_{i,k})}{\sum_{j \in \mathcal{N}_i} g(\theta_{i,j})}, \tag{21}$$

where $g(\cdot)$ is the function described in Figure 5 (b). The above weight function (21) can be seen as a linear interpolation of (20). We see from (21) that the weight on each edge is computable in a distributed manner; a node only needs to know the angular position of its neighbors. This design method does not require any knowledge of the network topology or centralized computation.

C. Numerical comparison

In this section, we present the numerical comparison of convergence rates among asymmetric design, symmetric optimal weights and weights chosen by the Metropolis-Hastings method. The symmetric optimal weights are obtained by using convex optimization method [9], [12]. The Metropolis-Hastings weights are picked by the following rule: $W_{i,j} = 1/|\mathcal{N}_i|$, where $\mathcal{N}_i$ denotes the number of node i's neighbors. The weights generated by this method are in general asymmetric. We plot the convergence rate R as a function of N, where N is the number of agents in the network. The amount of asymmetry used is $\varepsilon = 0.5$.

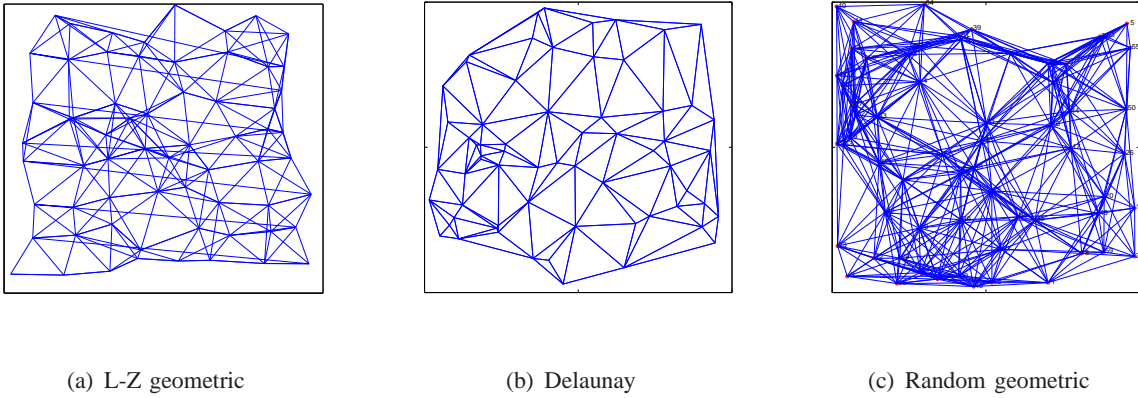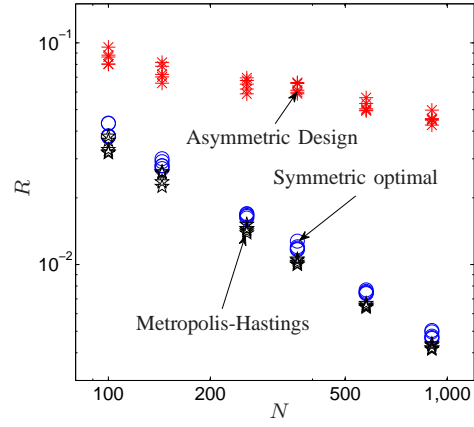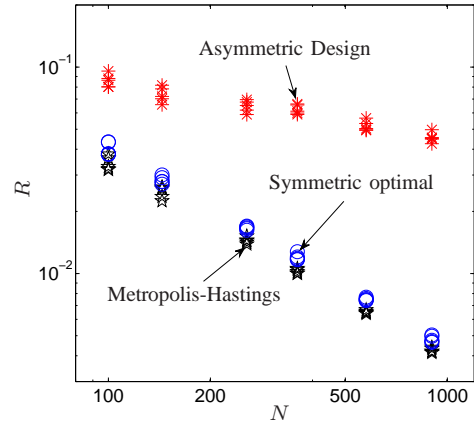(a) L-Z geometric   (b) Delaunay   (c) Random geometric

Fig. 6.   Examples of 2-D L-Z geometric, Delaunay and random geometric graphs.

We first consider a L-Z geometric graph [10], which is generated by perturbing the position of a square 2-D lattice ($N_1 = N_2 = \sqrt{N}$) with Gaussian random noise (zero mean and $1/(4\sqrt{N})$ standard deviation) and connect each nodes with the other nodes that are within $2/\sqrt{N}$ of radius neighborhood. Second, we consider a Delaunay graph [5], which is generated by placing N nodes on a 2-D unit square uniformly at random and connecting any two nodes if their corresponding Voronoi cells intersect, as long as their Euclidean distance is smaller than $1/3$. Finally, we consider a random geometric graphs [23], which is generated by placing N nodes on a 2-D unit square uniformly at random and connecting pairs of nodes that are within distance $3/\sqrt{N}$ of each other. Figure 6 gives examples of L-Z geometric graphs, Delaunay graphs and random geometric graphs.

(a) L-Z geometric graphs

(b) Delaunay graphs
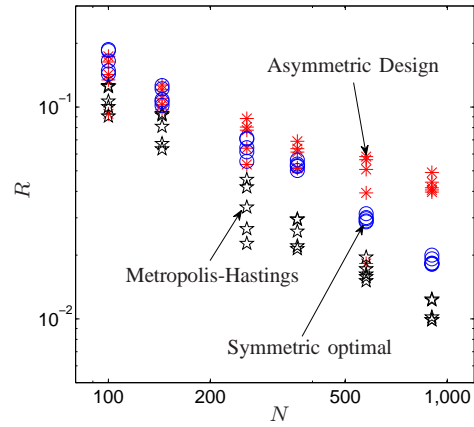
(c) Random geometric graphs

Fig. 7.   Comparison of convergence rates with proposed asymmetric weights, Metropolis-Hastings weights, and symmetric optimal. For each N, results from 5 sample graphs are plotted.

(a) L-Z geometric

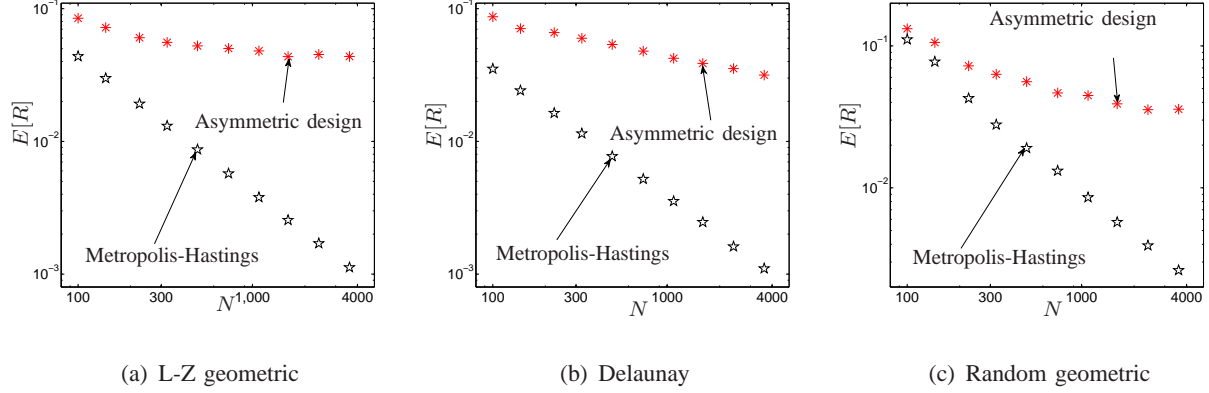(b) Delaunay

(c) Random geometric

Fig. 8. Mean of convergence rates of L-Z geometric, Delaunay and random geometric graphs with asymmetric design (AS) and Metropolis-Hastings (MH) for large N computed from 10 samples.

Figure 7 shows the comparison of convergence rates among asymmetric design, symmetric optimal and Metropolis-Hastings weights. For each N, the convergence rate of 10 samples of the graphs are plotted. We see from Figure 7 that for almost every sample in each of the three classes, the convergence rate with the asymmetric design is an order of magnitude larger than the others, especially when N is large. In addition, the convergence rates with symmetric optimal and Metropolis-Hastings methods are similar. Moreover, we observe from Figure 8 (a) and (c) that the slopes of the convergence rates with asymmetric design for L-Z geometric graphs and random geometric graphs are becoming progressively smaller with increasing N, which indicates that the convergence rate has a potential to be uniformly bounded below when N becomes arbitrarily large. The convergence rates with symmetric optimal weights are not included, since the numerical computations for large N are extremely expensive.

V. CONCLUSIONS AND FUTURE WORK

We studied the problem of how to design weights to increase the convergence rate of distributed consensus in networks with static topology. We proved that on lattice graphs, with proper choice of asymmetric weights, the convergence rate can be uniformly bounded away from zero. In addition, we propose a distributed weight design algorithm for 2-dimensional geometric graphs to improve the convergence rate, by using a continuum approximation. Numerical calculations show that the resulting convergence rate is substantially larger than that optimal symmetric

weights and Metropolis Hastings weights.

An important open question is a precise characterization of graphs for which theoretical guarantees on size-independent convergence rate can be provided with the proposed design. In addition, characterizing the asymptotic steady state value for more general graphs than lattices is also on-going work.

REFERENCES

[1] J. Tsitsiklis, "Problems in decentralized decision making and computation." Ph.D. dissertation, Massachusetts Institute of Technology, 1984.

[2] A. Jadbabaie, J. Lin, and A. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on Automatic Control*, vol. 48, no. 6, pp. 988–1001, 2003.

[3] R. Olfati-Saber, J. Fax, and R. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[4] W. Ren and R. Beard, *Distributed consensus in multi-vehicle cooperative control: theory and applications*. Springer Verlag, 2008.

[5] S. Martinez, J. Cortes, and F. Bullo, "Motion coordination with distributed information," *Control Systems Magazine, IEEE*, vol. 27, no. 4, pp. 75–88, 2007.

[6] A. Olshevsky and J. N. Tsitsiklis, "Convergence speed in distributed consensus and averaging," *SIAM J. Control and Optimization*, vol. 48, no. 1, pp. 33–55, 2009.

[7] T. Chan and L. Ning, "Fast convergence for consensus in dynamic networks," *Automata, Languages and Programming*, pp. 514–525, 2011.

[8] M. Cao, D. Spielman, and A. Morse, "A lower bound on convergence of a distributed network consensus algorithm," in *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. 44th IEEE Conference on*. IEEE, 2005, pp. 2356–2361.

[9] S. Boyd, P. Diaconis, and L. Xiao, "Fastest mixing markov chain on a graph," *SIAM review*, pp. 667–689, 2004.

[10] E. Lovisari and S. Zampieri, "Performance metrics in the average consensus problem: a tutorial," February 2012, annual Reviews in Control.

[11] R. Olfati-Saber, "Ultrafast consensus in small-world networks," in *American Control Conference, 2005. Proceedings of the 2005*. IEEE, 2005, pp. 2371–2378.

[12] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems & Control Letters*, vol. 53, no. 1, pp. 65–78, 2004.

[13] S. Boyd, "Convex optimization of graph Laplacian eigenvalues," in *Proceedings of the International Congress of Mathematicians*, vol. 3. Citeseer, 2006, pp. 1311–1319.

[14] Y. Cao and W. Ren, "Finite-time consensus for second-order multi-agent networks with inherent nonlinear dynamics under an undirected fixed graph," in *Proceedings of the 50th IEEE conference on Decision and Control*, 2011.

[15] L. Wang and F. Xiao, "Finite-time consensus problems for networks of dynamic agents," *Automatic Control, IEEE Transactions on*, vol. 55, no. 4, pp. 950–955, 2010.

[16] S. Sundaram and C. Hadjicostis, "Finite-time distributed consensus in graphs with time-invariant topologies," in *American Control Conference, 2007.*, 2007, pp. 711–716.

[17] H. Hao and P. Barooah, "Control of large 1D networks of double integrator agents: role of heterogeneity and asymmetry on stability margin," in *IEEE Conference on Decision and Control*, December 2010.

[18] P. Barooah, P. G. Mehta, and J. P. Hespanha, "Mistuning-based decentralized control of vehicular platoons for improved closed loop stability," *IEEE Transactions on Automatic Control*, vol. 54, no. 9, pp. 2100–2113, September 2009.

[19] P. G. Doyle and J. L. Snell, "Random walks and electric networks," Math. Assoc. of America, 1984.

[20] P. Barooah and J. P. Hespanha, "Error scaling laws for optimal estimation from relative measurements," *IEEE Transactions on Information Theory*, vol. 55, pp. 5661 – 5673, December 2009.

[21] H. Hao and P. Barooah, "Asymmetric control achieves size-independent stability margin in 1-d flocks," in *50th IEEE Conference on Decision and Control and European Control Conference*, December 2011, pp. 3458–3463.

[22] H. Hao, P. Barooah, and P. G. Mehta, "Stability margin scaling laws of distributed formation control as a function of network structure," *IEEE Transactions on Automatic Control*, vol. 56, pp. 923 – 929, April 2011.

[23] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2508–2530, 2006.

[24] W. Yueh and S. Cheng, "Explicit eigenvalues and inverses of tridiagonal toeplitz matrices with four perturbed corners," *The Australian & New Zealand Industrial and Applied Mathematics (Anziam) Journal*, vol. 49, no. 3, pp. 361–388, 2008.

[25] R. Haberman, *Elementary applied partial differential equations: with Fourier series and boundary value problems.* Prentice-Hall, 2003.

[26] L. Evans, *Partial Differential Equations: Second Edition (Graduate Studies in Mathematics).* American Mathematical Society, 2010.

APPENDIX

Proof of Lemma 1. The stochastic matrix $W^{(1)}$ has a simple eigenvalue $\lambda_1 = 1$. Following Theorem 3.1 of [24], the other eigenvalues of $W^{(1)}$ are given by

$$\lambda_\ell = 1 - a - c + 2\sqrt{ac}\cos\theta_\ell, \quad \ell \in \{2, \cdots, N\},$$

where θ_ℓ ($\theta \neq m\pi, m \in \mathbb{Z}$, $\mathbb{Z}$ being the set of integers) is the root of the following equation

$$2\sin(N\theta)cos(\theta) = (a+c)\sqrt{\frac{1}{ac}}\sin N\theta,$$

which implies

$$\sin(N\theta) = 0, \quad \text{or} \quad \cos\theta = \frac{(a+c)}{2}\sqrt{\frac{1}{ac}}.$$

Since $a > 0, c > 0$ and $a \neq c$, we have $\frac{(a+c)}{2}\sqrt{\frac{1}{ac}} > 1$, thus $\cos\theta \neq \frac{(a+c)}{2}\sqrt{\frac{1}{ac}}$. In addition, we have that $\theta \neq m\pi$, which yields

$$\theta_\ell = \frac{(\ell-1)\pi}{N}, \ \ell = \{2, \cdots, N\}. \tag{22}$$

We now obtain the eigenvalues of $W^{(1)}$, which is given by

$$\lambda_\ell = 1 - a - c + 2\sqrt{ac}\cos\frac{(\ell-1)\pi}{N}, \ \ell = \{2,\cdots,N\}.$$

Let $\pi = [\pi_1, \pi_2, \cdots, \pi_N]$ be the left Perron vector of $W^{(1)}$. From the definition of Perron vector, we have $\pi W^{(1)} = \pi$. Thanks to the special structure of the tridiagonal form of $W^{(1)}$, we can solve for π explicitly, which yields

$$\pi_i = (c/a)^{i-1}\pi_1, \tag{23}$$

where $i \in \{2, 3, \cdots, N\}$. In addition, we have $\pi_i > 0$ and $\sum_{i=1}^{N}\pi_i = 1$. Therefore,

$$1 = \sum_{i=1}^{N}\pi_i = \sum_{i=1}^{N}(c/a)^{i-1}\pi_1 \quad \Rightarrow \quad \pi_1 = \frac{1-c/a}{1-(c/a)^N}.$$

Substituting the above equation into (23), we complete the proof. ∎

Proof of Lemma 2. With the weights given in (8), it is straightforward - through a bit tedious - to show that the graph Laplacian $L^{(D)}$ associated with the D-dimensional lattice has the following form:

$$L^{(d)} = I_{N_d} \otimes L^{(d-1)} + L_d^{(1)} \otimes I_{N_1 N_2 \cdots N_{d-1}}, \quad 2 \leq d \leq D,$$

where $L^{(1)} = L_1^{(1)}$ and $L_d^{(1)} = 1 - W_d^{(1)}$ is the Laplacian matrix of dimension $N_d \times N_d$, which is given by

$$L_d^{(1)} = \begin{bmatrix} c_d & -c_d & & & \\ -a_d & a_d + c_d & -c_d & & \\ & \ddots & \ddots & \ddots & \\ & & -a_d & a_d + c_d & -c_d \\ & & & -a_d & a_d \end{bmatrix}. \tag{24}$$

Since a D-dimensional lattice is the Cartesian product graph of D 1-dimensional lattices, the eigenvalues of the graph Laplacian matrix $L^{(D)}$ are sum of the eigenvalues of the D 1-dimensional Laplacian matrix $L_d^{(1)}$. Thus, we have

$$\mu_{\ell_1,\ldots,\ell_D}(L^{(D)}) = \sum_{d=1}^{D}\mu_{\ell_d}(L_d^{(1)}).$$

In addition, we have that $W^{(D)} = I_N - L^{(D)}$ and $W_d^{(1)} = I_{N_d} - L_d^{(1)}$, thus the eigenvalues $\lambda_{\vec{\ell}}$ of $W^{(D)}$ are given by

$$\lambda_{\vec{\ell}}\left(W^{(D)}\right) = 1 - \mu_{\vec{\ell}}\left(L^{(D)}\right) = 1 - \sum_{d=1}^{D} \mu_{\ell_d}(L_d^{(1)})$$

$$= 1 - \sum_{d=1}^{D}(1 - \lambda_{\ell_d}(W_d^{(1)})).$$

To see $\pi = \pi_D^{(1)} \otimes \pi_{D-1}^{(1)} \otimes \cdots \otimes \pi_1^{(1)}$ is the left Perron vector of $W^{(D)}$, we first notice that

$$\pi_d^{(1)} W_d^{(1)} = \pi_d^{(1)}, \quad \pi_d^{(1)} L_d^{(1)} = 0,$$

where $d \in \{1, \cdots, D\}$. The rest of the proof follows by straightforward induction method, we omit the proof due to space limit. ∎

Proof of Lemma 3. Multiply both sides of (14) by $2N^2 e^{2\varepsilon Nx}$, we obtain the standard Sturm-Liouville eigenvalue problem

$$\frac{d}{dx}\left(e^{2\varepsilon Nx}\frac{d\phi(x)}{dx}\right) + 2N^2 \mu e^{2\varepsilon Nx}\phi(x) = 0. \tag{25}$$

According to Sturm-Liouville Theory, all the eigenvalues are real, see [25], [26]. To solve the Sturm-Liouville eigenvalue problem (14)-(15), we assume solution of the form, $\phi(x) = e^{rx}$, then we obtain the following equation

$$r^2 + 2\varepsilon Nr + 2\mu N^2 = 0,$$

$$\Rightarrow r = N(-\varepsilon \pm \sqrt{\varepsilon^2 - 2\mu}). \tag{26}$$

Depending on the discriminant in the above equation, there are three cases to analyze:

1) $\mu < \varepsilon^2/2$, then the eigenfunction $\phi(x)$ has the following form $\phi(x) = c_1 e^{N(-\varepsilon+\sqrt{\varepsilon^2-2\mu})x} + c_2 e^{N(-\varepsilon-\sqrt{\varepsilon^2-2\mu})x}$, where c_1, c_2 are some constants. Applying the boundary condition (15), it's straightforward to see that, for non-trivial eigenfunctions $\phi(x)$ to exit, the following equation must be satisfied

$$\frac{-\varepsilon + \sqrt{\varepsilon^2 - 2\mu}}{\varepsilon + \sqrt{\varepsilon^2 - 2\mu}} = e^{2N\sqrt{\varepsilon^2-2\mu}}\frac{-\varepsilon + \sqrt{\varepsilon^2 - 2\mu}}{\varepsilon + \sqrt{\varepsilon^2 - 2\mu}}.$$

Thus, we have $\mu = 0$.

2) $\mu = \varepsilon^2/2$, then the eigenfunction $\phi(x)$ has the following form

$$\phi(x) = c_1 e^{-\varepsilon Nx} + c_2 x e^{-\varepsilon Nx}.$$

Applying the boundary condition (15) again, it's straightforward to see that there is no eigenvalue for this case.

3) $\mu > \varepsilon^2/2$, then the eigenfunction has the following form $\phi(x) = e^{-\varepsilon N x}(c_1 \cos(N\sqrt{2\mu - \varepsilon^2}x) + c_2 \sin(N\sqrt{2\mu - \varepsilon^2}x)$. Applying the boundary condition (15), for non-trivial eigenfunctions to exit, the eigenvalues μ must satisfy $\mu = \frac{\varepsilon^2}{2} + \frac{\ell^2 \pi^2}{2N^2}$, where $\ell = 1, 2, \cdots$.

Combining the above three cases, the eigenvalues of the Sturm-Liouville operator are $\mu \in \{0, \frac{\varepsilon^2}{2} + \frac{\ell^2 \pi^2}{2N^2}\}$, where $\ell \in \{1, 2, \cdots\}$. The second smallest eigenvalue $\mu_2(\mathcal{L})$ of the Strum-Liouville operator $\mathcal{L}$ is then given by

$$\mu_2(\mathcal{L}) = \frac{\varepsilon^2}{2} + \frac{\pi^2}{2N^2} \geq \frac{\varepsilon^2}{2},$$

which is a constant that is bounded away from 0. ■

Proof of Theorem 2. According to Lemma 1, the eigenvalues of $W_d^{(1)}$ are given by:

$$\lambda_1(W_d^{(1)}) = 1,$$

$$\lambda_\ell(W_d^{(1)}) = 1 - a_d - c_d + 2\sqrt{a_d c_d} \cos \frac{(\ell_d - 1)\pi}{N_d}.$$

From Lemma 2, the second largest eigenvalue $\lambda_2(W^{(D)})$ and the smallest eigenvalue $\lambda_N(W^{(D)})$ of $W^{(D)}$ are given by

$$\lambda_2(W^{(D)}) = 1 - \max_{d \in \{1, \cdots, D\}} (1 - \lambda_2(W_d^{(1)}))$$

$$\leq 1 - \max_{d \in \{1, \cdots, D\}} (a_d + c_d - 2\sqrt{a_d c_d}), \tag{27}$$

$$\lambda_N(W^{(D)}) = 1 - \sum_{d=1}^{D} (1 - \lambda_{N_d}(W_d^{(1)}))$$

$$= 1 - \sum_{d=1}^{D} (a_d + c_d - 2\sqrt{a_d c_d} \cos \frac{(N_d - 1)\pi}{N_d})$$

$$\geq 1 - \sum_{d=1}^{D} (a_d + c_d - 2\sqrt{a_d c_d}). \tag{28}$$

Recall that $R = \min\{1 - \lambda_2, 1 + \lambda_N\}$. In addition, a_d, c_d are fixed constants and satisfy $a_d \neq c_d$, $\sum_{d=1}^{D} a_d + c_d \leq 1$, therefore the lower bounds of $1 - \lambda_2(W^{(D)})$ and $1 + \lambda_N(W^{(D)})$ are fixed positive constants. We then have that the convergence rate of $W^{(D)}$ satisfy $R = 1 - \rho(W^{(D)}) \geq c_0$, where c_0 is a constant independent of N. ■

Proof of Theorem 3. By the method of separation of variables [25], [26], the eigenvalues of the Sturm-Liouville operator $\mathcal{L}^{(D)}$ is given by

$$\mu(\mathcal{L}^{(D)}) = \sum_{d=1}^{D} \mu(\mathcal{L}_d^{(1)}), \qquad (29)$$

where $\mathcal{L}_d^{(1)}$ is the 1-dimensional Sturm-Liouville operator given by

$$\mathcal{L}_d^{(1)} = -\frac{1}{2DN_d^2}\frac{d^2}{dx_d^2} - \frac{\varepsilon}{DN_d}\frac{d}{dx_d},$$

with Neumann boundary conditions. Following Lemma 3, we have that the smallest eigenvalue of $\mathcal{L}_d^{(1)}$ is 0 and the second smallest eigenvalue of $\mathcal{L}_d^{(1)}$ is bounded below by $\mathcal{L}_d^{(1)} \geq \varepsilon^2/2D$. Therefore, we have from (29) that the second smallest eigenvalue is

$$\mu_2(\mathcal{L}^{(D)}) = \min_d\{\mu_2(\mathcal{L}^{(d)})\} \geq \frac{\varepsilon^2}{2D}. \qquad \blacksquare$$